(12) EUROPEAN PATENT APPLICATION

(72) Inventor : Wong, Ping W.
1443 Knowlton Drive
Sunnyvale, CA 94087 (US)

(74) Representative : Powell, Stephen David et al
WILLIAMS, POWELL & ASSOCIATES
34 Tavistock Street
London WC2E 7PB (GB)

(54) Inverse halftoning using linear filtering and statistical smoothing.

(57) A method and apparatus for converting $T(\cdot)$ a binary image $y_{m,n}$ into a gray scale image $T(y_{m,n})$ includes a pipeline having a number of nearly identical stages $T(\cdot)$. At each stage, modifications occur to provide the gray scale image by following a sequence of steps including : (1) determining a sample mean value representing an averaging of pixel values within a neighborhood containing a selected pixel ; (2) determining a weighted variation value representing pixel value variations within the neighborhood ; and (3) adjusting the pixel value of the selected pixel toward the sample mean value if the difference between the pixel value and the sample mean value exceeds the weighted variation value. The adjustment to the pixel value is in the direction of the sample mean value. Preferably, the weighting of the variation value is different at each stage. At each stage along the pipeline, low pass filtering (18) is provided to remove high frequency noise generated by the process.
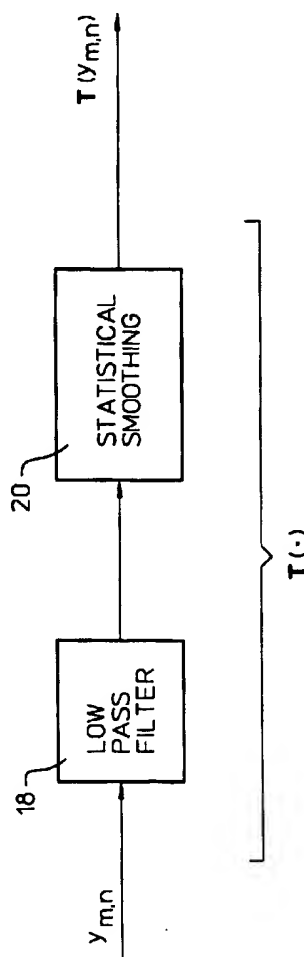
FIG. 2

The present invention relates to a method and apparatus for constructing a high quality gray scale image from a binary image.

Background Art

5

"Halftoning" is the process of constructing a binary image, i.e. a bi-level image, from a gray scale image, i.e. a multi-level image. The constructed binary image can then be displayed by means of binary devices, such as newspaper printers and laser printers. An n-bit gray scale image is converted into an one-bit binary image perceived to contain a continuous tone. The gray scale image may be passed through a sequence of operators

10   that assign pixel values of 1 or 0 for each pixel of the resulting binary image.

"Inverse halftoning" is the process of reconverting the binary image into an approximation of the original gray scale image. Inverse halftoning can be applied to a wide variety of binary image processing problems, such as scaling, tone correction, interchanging between halftone methods, facsimile image processing, and image compression. For example, a gray scale image may be constructed from a binary image, whereafter

15   the gray scale image undergoes a processing operation and is finally rehalftoned.

A binary image is a two-dimensional array of pixels having pixel values limited to two levels. On the other hand, a gray scale image is a two-dimensional light intensity function $x(i,j)$, where i and j denote discrete pixel coordinates and where the pixel value at any pixel x is proportional to the gray level of the image at that point. For example, the gray scale image may be an eight-bit scale having 256 possible gray levels.

20   Halftoning methods that use algorithms can generally be classified into two different categories -- ordered dither and error diffusion. Ordered dither can be described at a basic level as a procedure of converting the gray scale image by thresholding the gray input with a periodically repeated threshold matrix. Error diffusion can be described as "spreading out" the error between an input gray-scale pixel and its binary output over a small area of the binary image. A weighted combination of output errors from previously processed pixels with

25   respect to the scanning strategy is added to the input pixel, and the sum is thresholded to produce the binary output. Thus, error diffusion provides a local average within any small area of the binary image to approximate the gray level in the corresponding area of the gray scale image.

The two types of halftoning produce binary images that are substantially different both with respect to structure and characteristics, e.g., the frequency spectra. Consequently, inverse halftoning methods designed

30   for converting one type of halftone image generally do not work well for converting the other type. With regard to inverse halftoning of dithered images, one technique is to use a "neighborhood approach" using adaptive run-lengths of 1's and 0's (ABRL). This performs particularly well in a three-level cascade algorithm comprised of ABRL, statistical smoothing, and impulse removal. Miceli et al., "Inverse Halftoning," Journal of Electronic Imaging, column, vol. 1, pages 143-151, April 1992.

35   With respect to inverse halftoning of error diffused images, the use of look-up tables has recently been suggested. Ting et al., "Error Diffused Image Compression Using a Halftone-to-Gray Scale Decoder and Predictive Pruned Tree-Structured Vector Quantization," submitted to IEEE Transactions on Image Processing, 1992. Ting et al. describe using a small window which is slid over the error diffused image. The content of the binary pixels in the window serves as an address to a look-up table. A gray level value is then retrieved as the

40   constructed gray level of the center pixel in the window. Thus, the inverse halftoning procedure can be interpreted as a decoding operation, where the "decoder" is the look-up table that associates a particular gray level with a particular bit pattern. The look-up table is constructed by a training algorithm using a collection of test images in the same manner as designing vector quantizers. While this method of inverse halftoning of an error diffused image works reasonably well, the training of the look-up table is time consuming. Moreover, because

45   error diffusion produces a binary image in which the local average within any small area of the image approximates the gray level of the corresponding small area of the original gray scale image, one can obtain very different bit patterns corresponding to the same local gray level by simply shifting a halftoning window a small amount in any direction. The different bit patterns will result in different addressing to the look-up table during inverse halftoning. On the other hand, one can generate very different binary images from the same gray scale

50   image by merely changing the initial condition in the error diffusion process. This means that the correlation between a specific bit pattern with the original local gray level is not very high, and hence the performance of the look-up table approach is limited.

Low pass filtering is the conventional way of reconstructing gray scale images from binary images. However, while the error diffusion process can be interpreted as one that injects noise into the gray scale image

55   primarily at the high frequency range, there are several reasons why low pass filtering alone does not produce inverse halftoned images of sufficient quality. Firstly, low pass filtering cannot remove all of the noise (i.e. error) purposefully introduced by the error diffusion process, since there are noise components at the low frequency range. Secondly, there may have been high frequency components in the original gray scale image,

2

so that removal of noise at the high frequency range will also remove these desirable high frequency components. Low pass filters are not so selective as to unambiguously determine whether a high frequency component originated from the gray scale image or from the error diffusion process. An example of low pass filtering is one which uses a filter at one-half of the maximum possible bandwidth of the digital image, i.e., the cut-off frequency is at $f_s/4$ where f_s is the sampling frequency. A significant amount of residue noise resulting from the error diffusion process will still remain at the output image. The residue noise cannot be removed by further low pass filtering unless the cut-off frequency of the filter is lowered significantly. However, lowering the cut-off frequency further would result in the removal of desired components, thereby generating an overly blurred image.

It is an object of the present invention to provide a method and apparatus for converting a binary image into a high quality gray scale image.

Summary of the Invention

The above object has been met by a gray scale image-constructing method and apparatus which utilize non-linear statistical smoothing to distinguish between desired and undesired components of an image to be converted. In a preferred embodiment, errors generated during either a halftoning process or an inverse halftoning process are removed and a high quality gray scale image is constructed by using a pipeline of stages, with each stage having both low pass filtering and non-linear statistical smoothing.

The purpose of the low pass filter at each stage of the pipeline is to remove undesired high frequency components generated by the non-linear smoothing process. The low pass filters are preferably linear filters. Each of the filters may be a halfband filter, or the filters may have different cut-off frequencies with progression along the pipeline.

The non-linear statistical smoothing is performed on a pixel-by-pixel basis and includes selecting a neighborhood for each pixel within a binary image to be converted. A first value that represents an averaging of pixel values within the neighborhood is determined. Also determined is a second value that characterizes the amount of local variation among pixel values within the neighborhood. The pixel value of a selected pixel, typically the center pixel in the neighborhood, is adjusted only if the difference between the pixel value and the first value exceeds a weighted multiple of the second value. The adjustment is toward the first value. This is repeated one pixel at a time until each pixel has been considered, thereby forming a gray scale image. At the first stage, the gray scale image is then passed to the second stage for repeating the low pass filtering and non-linear statistical smoothing.

The center pixel in the selected neighborhood has the coordinates m,n. The size of the neighborhood ($R_{m,n}$) is not critical, but should be reasonably small. By way of example, the neighborhood may have a size of 3 by 3. The first value ($\mu_{m,n}$) is the sample mean of the pixel values in $R_{m,n}$. The second value ($v_{m,n}$) characterizes locally the amount of variations among the pixels within $R_{m,n}$. The pixel value ($x_{m,n}^{old}$) of the center pixel is compared with the thresholds $\mu_{m,n} + \gamma v_{m,n}$ and $\mu_{m,n} - \gamma v_{m,n}$, where γ is a weighted multiple that is greater than or equal to 0 and is adjusted on a per-stage basis to obtain optimal performance. The value of the center pixel is modified to provide an adjusted pixel value ($x_{m,n}^{new}$) satisfying the condition:

$$
x_{m,n}^{new} = \begin{cases}
\mu_{m,n} + \gamma v_{m,n} & \text{if } x_{m,n}^{old} > \mu_{m,n} + \gamma v_{m,n} \\
\mu_{m,n} - \gamma v_{m,n} & \text{if } x_{m,n}^{old} < \mu_{m,n} - \gamma v_{m,n} \\
x_{m,n}^{old} & \text{otherwise.}
\end{cases}
$$

A window that defines the neighborhood "slides" across the input image in a pixel-by-pixel manner so as to consider each pixel in the input image. The output image is therefore a multi-level image. The output of each stage in the pipeline is the input to the subsequent stage. The signal that exits from the pipeline may be displayed, or printed using a multi-level printer, or may be caused to undergo a processing operation such as image compression.

While not critical, an example of a calculation of $v_{m,n}$ for measuring the local variation is the rth power norm centered at the sample mean. Thus,

$$v_{m,n} = \left[\frac{1}{\|R_{m,n}\|} \sum_{(i,j) \in R_{m,n}} \left| x_{i,j}^{old} - \mu_{m,n} \right|^r \right]^{1/r},$$

where

$$\mu_{m,n} = \frac{1}{\|R_{m,n}\|} \sum_{(i,j) \in R_{m,n}} x_{i,j}^{old}$$

and where $\|R_{m,n}\|$ denotes the number of pixels in the neighborhood. In the special case where $r = 2$, $v_{m,n}$ is identical to the sample standard deviation of the pixels in $R_{m,n}$.

An advantage of the present invention is that non-linear smoothing discriminates in the removal of high frequency components. Noise components introduced by error diffusion are removed, so that the resulting image is not blurred to the extent that is typical of using low pass filtering with a low cut-off frequency. A pixel value is modified only when the pixel value is statistically far away from the sample average within the neighborhood. Both $\mu_{m,n}$ and $v_{m,n}$ vary with the position of the window, so that statistical smoothing is performed strictly according to local statistics. When the window falls in a location that contains an edge of an image, the variation measure $v_{m,n}$ tends to be larger than in the case in which there is not an edge within the window. Consequently, the pixel value of a pixel is less likely to be modified when the pixel is near an edge. As a result, the smoothing operation is performed without overly blurring the image. Another advantage is that the invention may be applied equally well whether the binary-to-gray scale conversion is a conversion from a bi-tone to a continuous tone image or a conversion from a bi-level color approach to a multi-level color approach. For example, three parallel pipelines may be used to convert three bi-level color signals to three multi-level color signals to drive the different color guns of a computer monitor, thereby providing a greater color spectrum of an image at the monitor.

Exemplary embodiments will now be described with reference to the drawings in which:

Fig. 1 is a block diagram schematically showing a prior art error diffusion halftoning system.

Fig. 2 is a block diagram schematically showing one stage of an inverse halftoning apparatus in accordance with the present invention.

Fig. 3 is a block diagram schematically showing a pipeline inverse halftoning apparatus in accordance with the present invention.

Best Mode for Carrying Out the Invention

With reference to Fig. 1, a generic error diffusion apparatus is shown as including a binary quantizer $Q(\cdot)$ 10 and a digital filter 12. The input $x_{m,n}$ is a gray scale image having a dynamic range L, i.e., $x_{m,n} \in [O,L]$ for all m and n. The output $b_{m,n}$ is a halftoned version of $x_{m,n}$, wherein $b_{m,n} \in \{0,1\}$. In the absence of a signal from the digital filter 12, the input passes unmodified through a summer 14 to the quantizer.

The threshold of the binary quantizer 10 is typically set at $L/2$. The output of the quantizer 10 is combined with the input at a second summer 16 to generate the signal $e_{m,n}$. The digital filter is a linear filter in a negative feedback loop. Thus, the filter 12 operates on the quantization error generated by the one-bit quantizer 10.

The error diffusion apparatus of Fig. 1 operates much in the manner of a predictive coder. The digital filter 12 has finite-extent input and output masks. Error diffusion "spreads out" the error between an input gray-scale pixel $(x_{m,n})$ and its binary equivalent $(b_{m,n})$ over a small area of the halftoned image. The digital filter 12 in the feedback loop is "causal" with respect to the scanning direction of a window along the gray scale image, so that only "past" errors are diffused.

An inverse halftoning procedure attempts to reproduce the gray scale image which was initially inputted to the apparatus of Fig. 1. With reference to Fig. 2, a stage of a pipeline apparatus for inverse halftoning is shown. The stage includes a low pass filter 18 and a statistical smoothing device 20. The stage is referred to by the designation $T(\cdot)$.

As shown in Fig. 3, a pipeline for inverse halftoning of error diffused images includes a number (k) of the stages of Fig. 2.

4

Returning to Fig. 2, the purpose of the low pass filter at each stage is to remove the potentially unwanted high frequency components that are generated by the non-linear smoothing process at 20. The low pass filter may be a halfband filter to avoid overly blurring the reconstructed image $T(y_{m,n})$. Alternatively, the filters may vary along the pipeline with respect to the cut-off frequency, i.e., the k filters are not necessarily identical. Low pass filters are conventionally used in the art of inverse halftoning, so that the configuration of the filters is readily understood.

Non-linear statistical smoothing 20 is preferably performed by means of the following steps. For each pixel location (m,n), a small neighborhood $R_{m,n}$ that is centered by the pixel location is defined. The sample mean $\mu_{m,n}$ of the pixels in the neighborhood is computed. Also determined is a measure $v_{m,n}$ that characterizes locally the amount of variation among the pixels within the neighborhood. The value of the center pixel is compared with the threshold $\mu_{m,n} + \gamma v_{m,n}$ and $\mu_{m,n} - \gamma v_{m,n}$. The multiple preferably is varied along the various stages 22 and 24 of Fig. 3. The weighted multiple should be greater than or equal to 0 and is preferably less than or equal to 1.

The value of the center pixel $(x_{m,n})$ is modified according to

$$
x_{m,n}^{new} = \begin{cases} \mu_{m,n} + \gamma v_{m,n} & \text{if } x_{m,n}^{old} > \mu_{m,n} + \gamma v_{m,n} \\ \mu_{m,n} - \gamma v_{m,n} & \text{if } x_{m,n}^{old} < \mu_{m,n} - \gamma v_{m,n} \\ x_{m,n}^{old} & \text{otherwise.} \end{cases}
$$

The superscripts "old" and "new" refer to the value of the center pixel before and after, respectively, each statistical smoothing process along the pipeline of Fig. 3. The operation is repeated on a pixel-by-pixel basis over the entire image, wherein the window that defines neighborhood "slides" across the image.

An example of $v_{m,n}$ for measuring the local variation is the rth power norm centered at the sample mean. Thus,

$$
v_{m,n} = \left[\frac{1}{\|R_{m,n}\|} \sum_{(i,j) \in R_{m,n}} \left| x_{i,j}^{old} - \mu_{m,n} \right|^r \right]^{1/r},
$$

where

$$
\mu_{m,n} = \frac{1}{\|R_{m,n}\|} \sum_{(i,j) \in R_{m,n}} x_{i,j}^{old}
$$

and where $\|R_{m,n}\|$ denotes the number of pixels in the neighborhood $R_{m,n}$. In the special case where R = 2, $v_{m,n}$ is identical to the sample standard deviation of the pixels in the neighborhood. While other determinations of the local variation measure are possible, the above-identified determination of $v_{m,n}$ has been shown to function well in providing high quality inverse halftoning.

The operation of each non-linear statistical smoothing is such that a particular pixel value is changed only when it is statistically far away from the sample mean of its neighbors. If the value of the center pixel in the neighborhood differs from $\mu_{m,n}$ by more than $v_{m,n}$, the pixel value of the center pixel is adjusted towards the sample mean $\mu_{m,n}$. If the difference between the original pixel value is sufficiently close to the sample mean, the pixel value remains unchanged.

Both $\mu_{m,n}$ and $v_{m,n}$ vary with the position of the window that defines the neighborhood. Therefore, the statistical smoothing operation is performed strictly according to local statistics. When the window falls in an area that contains an edge of the image, the sample variant $v_{m,n}$ of the defined neighborhood $R_{m,n}$ tends to be larger than in the case in which the window does not contain an edge. It follows that the pixel value of the center pixel

5

of a neighborhood is less likely to be adjusted when it is near the edge. Consequently, the nonlinear statistical smoothing operation provides a gray scale image without overly blurring the image at its edges.

As can be seen from the formula above for determining $x_{m,n}^{new}$, a smaller value of γ implies that it is more likely that the pixel value of the center pixel will exceed either the upper or lower threshold. Adjustments of the pixel values will therefore occur more frequently throughout the image. For this reason, it is expected that a smoother output image results in use with a smaller γ. In experiments in which the window had a size of 3 by 3 and r = 2 in the determination of $v_{m,n}$, the best performance was obtained by setting the initial weighted multiple (γ) to 0.25 and then linearly increasing the weighted multiple along four stages 20-22 to reach a final value of 1.0. It was also found that varying the cut-off frequency along the low pass filters 18 of the various stages did not improve the reconstructed image.

As shown in Fig. 3, a final low pass filter 26 is provided to remove the unwanted high frequency components that may be generated at the final stage 24 of the pipeline. The resulting gray scale image may be presented for display at a monitor 28, or may be printed using a multi-level printer or may be caused to undergo some other form of image processing, such as compression.

The present invention has been described as one in which it is the center pixel in a defined neighborhood that is considered for adjustment. While this is preferred, it is possible to select some other pixel within the neighborhood. Moreover, each pixel within a neighborhood defined by a window may be considered for modification before relocation of the window. However, a loss of quality may result.

The invention is not restricted to conversion of a binary image into a continuous tone image. The invention applies equally well to a conversion from an image that is binary with respect to color into an image that is "gray scale," i.e. multi-level, with respect to color. That is, a "gray scale image" is defined herein as a multi-level image with respect to image-viewable factors such as tone or color. For example, a color monitor having a separate color gun for each of three primary colors may be driven by an apparatus in which three separate pipelines convert bi-level color images to gray scale color images for combination at the display screen.

Claims

1. A method for converting a binary image into a gray scale image comprising:
    (a) selecting a neighborhood of pixels in the binary image $Y_{m,n}$, each pixel having a pixel value at one of two levels, said neighborhood including a selected pixel;
    (b) determining (20) a first value responsive to an averaging of pixel values of said pixels within said neighborhood;
    (c) determining (20) a second value responsive to a characterization of variation of pixel values among said pixels within said neighborhood;
    (d) increasing (20) the pixel value of said selected pixel if said pixel value is less than the difference determined by subtracting said second value from said first value;
    (e) decreasing (20) said pixel value of said selected pixel if said pixel value is greater than the sum of said first and second values; and
    (f) repeating steps (a)-(e) for other pixels of said binary image such that a gray scale image $T(y_{m,n})$ is formed.

2. The method of claim 1 wherein determining (20) the second value includes mathematically weighting said characterization of said variation of pixel values, thereby providing said second value.

3. The method of claim 2 further comprising repeating $T(\cdot)$ steps (a)-(f) for said gray scale image $T(y_{m,n})$ formed by said steps, wherein repeating steps (a)-(f) is performed with a different mathematical weighting in determining said second value.

4. The method of claim 1 further comprising presenting said gray scale image $T(y_{m,n})$ at a visual media (28).

5. The method of claims 1, 2 and 4 wherein said selected pixel ($x_{m,n}$) is at coordinates m,n in a two-dimensional array of pixels and said first and second values are $\mu_{m,n}$ and $\gamma v_{m,n}$, respectively, said steps providing an adjusted pixel value ($x_{m,n}^{new}$) wherein

$$
x_{m,n}^{new} = \begin{cases} \mu_{m,n} + \gamma v_{m,n} & \text{if } x_{m,n}^{old} > \mu_{m,n} + \gamma v_{m,n} \\ \mu_{m,n} - \gamma v_{m,n} & \text{if } x_{m,n}^{old} < \mu_{m,n} - \gamma v_{m,n} \\ x_{m,n}^{old} & \text{otherwise.} \end{cases}
$$

6. The method of claim 5 wherein said step of determining the second value ($\gamma v_{m,n}$) includes multiplying a weighted value (γ) times the local variation ($v_{m,n}$) of pixel values within said neighborhood ($R_{m,n}$) of pixels such that at the rth power norm
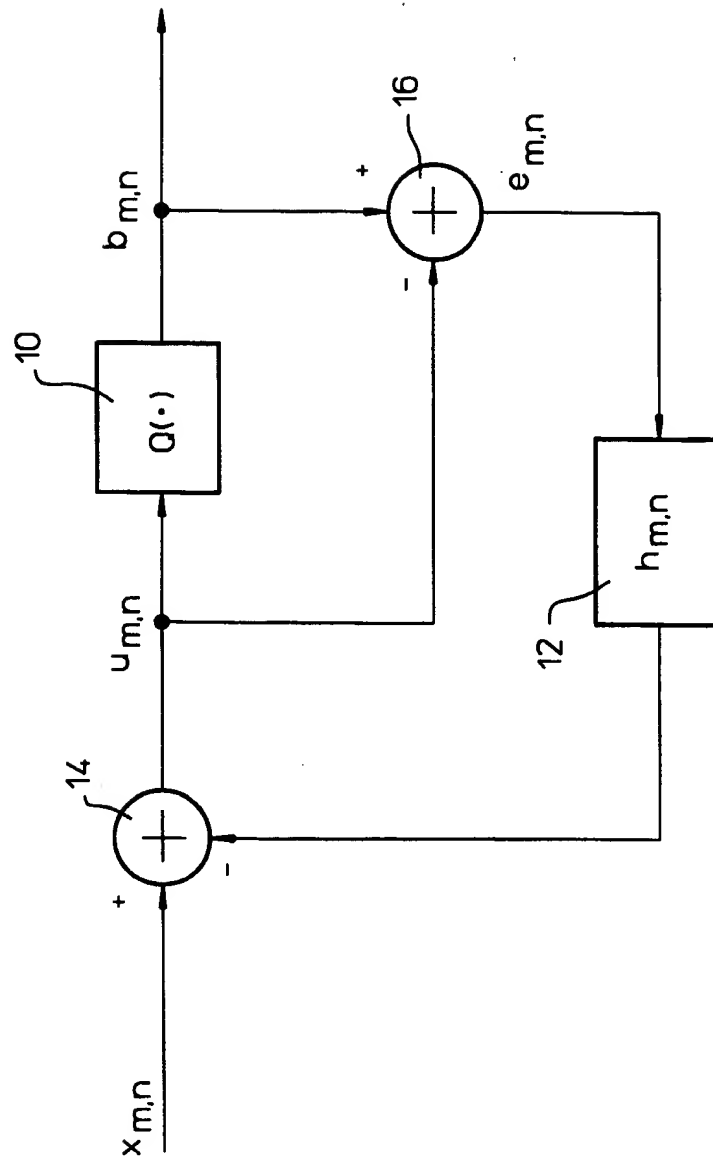
$$
v_{m,n} = \left[\frac{1}{\|R_{m,n}\|} \sum_{(i,j) \in R_{m,n}} \left| x_{i,j}^{old} - \mu_{m,n} \right|^r \right]^{1/r},
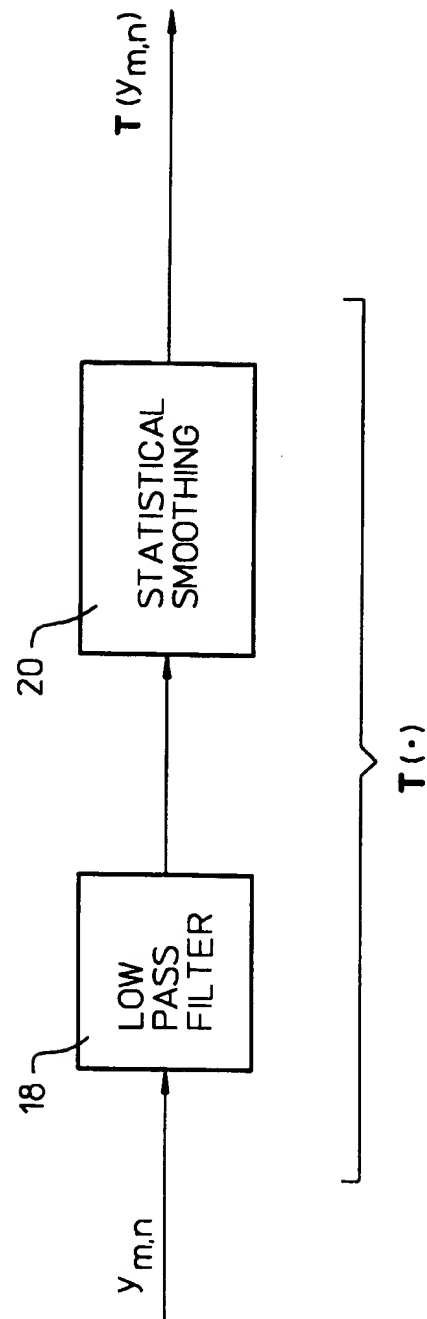$$

where

$$
\mu_{m,n} = \frac{1}{\|R_{m,n}\|} \sum_{(i,j) \in R_{m,n}} x_{i,j}^{old}
$$

and where $\|R_{m,n}\|$ denotes the number of pixels in the neighborhood.

7. The method of claims 1, 2 and 4 further comprising low pass filtering said gray scale image $T(y_{m,n})$ formed by steps (a)-(f).

8. The method of claims 1, 2 and 4 wherein said binary image $y_{m,n}$ is a bi-level color image and wherein the formed gray scale image $T(y_{m,n})$ is a multi-level color image.

$x_{m,n}$

$u_{m,n}$

$b_{m,n}$

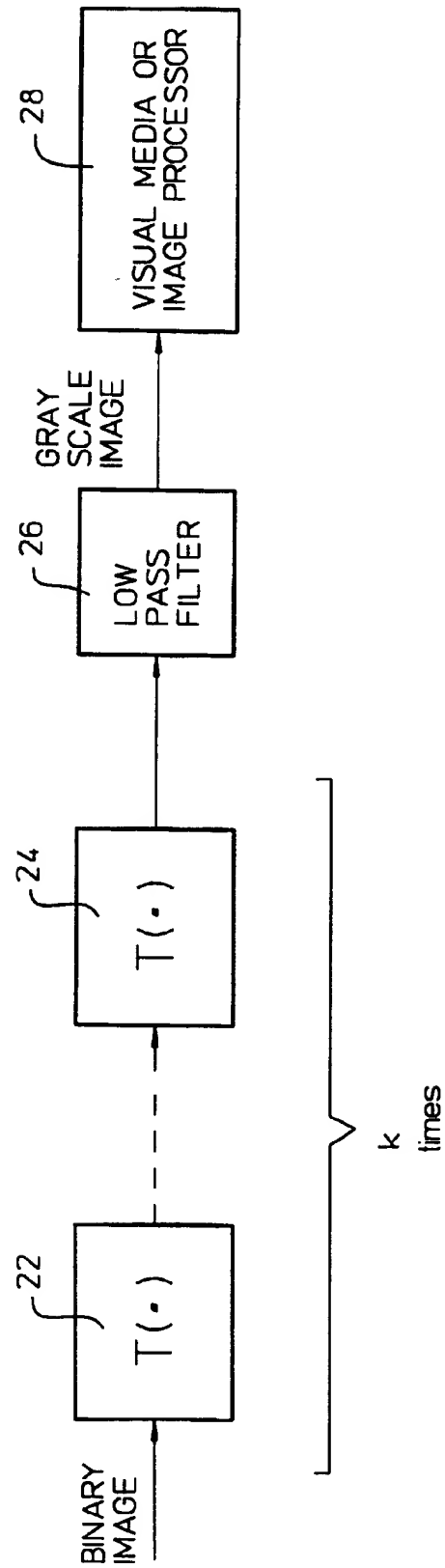$e_{m,n}$

$Q(\cdot)$ 10

14

16

$h_{m,n}$ 12

FIG. 1 (PRIOR ART)

FIG. 2

BINARY IMAGE → T(·) [22] — — — → T(·) [24] → LOW PASS FILTER [26] → GRAY SCALE IMAGE → VISUAL MEDIA OR IMAGE PROCESSOR [28]

k times

FIG. 3

10

European Patent Office

EUROPEAN SEARCH REPORT

Application Number

EP 94 30 2654

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.5)
A,P	EP-A-0 562 813 (XEROX CORPORATION) * abstract; figures * ---	1	H04N1/40
A	HUMAN VISION, VISUAL PROCESSING, AND DIGITAL DISPLAY IV, PROCEEDINGS OF THE SPIE, vol.1913, February 1993, SAN JOSE, CA, USA pages 310 - 324 HEIN S ; ZAKHOR A 'Reconstruction of continuous tone images from their error diffused halftone version' * the whole document * -----	1	

TECHNICAL FIELDS SEARCHED (Int.Cl.5)

H04N

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	28 July 1994	Isa, S

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

11